

DOCUMENT RESUME

ED 151 413

TM 006 936

AUTHOR Glick, Oren
TITLE Evaluating Innovative Programs: Paper Prepared for Distribution in Conjunction with the 1976 AERA Symposium Presentation: Internal Evaluation in the Experimental Schools Projects.
PUB DATE Apr 76
NOTE 53p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)
EDRS PRICE MF-\$0.83 HC-\$3.50 Plus Postage.
DESCRIPTORS Academic Achievement; Cost Effectiveness; Curriculum Development; Data Processing; Decision Making; *Educational Alternatives; *Educational Innovation; Elementary Secondary Education; *Evaluation Methods; Evaluators; Experimental Schools; Individualized Instruction; Information Utilization; Models; *Program Evaluation; Research Problems; Student Attitudes; Teacher Attitudes; *Testing Programs
IDENTIFIERS Franklin Pierce School District WA; Washington (Tacoma)

ABSTRACT

The Franklin Pierce School District, a semi-rural, semi-urban district located south of Tacoma, Washington, was one of the Experimental School Projects (ESP) initially funded by the National Institute of Education. The Franklin Pierce ESP provides an optimal learning environment for each student, with special emphasis placed upon transient, gifted, and high risk students. A variety of proven educational system and curriculum models are implemented in an integrated fashion. The ESP evaluation plan identified three major goals: (1) to provide the Superintendent and the school board with accurate information on the progress and effectiveness of the project's programs, (2) to set up and operate a program monitoring process that provides teaching staff and district administrators with information needed in decision making, and (3) to accumulate, and periodically report to federal sources, appropriate data related to the progress of the program in fulfilling its objectives.
(Author/MV)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Paper Prepared for Distribution in
Conjunction With the 1976 AERA Symposium Presentation:
Internal Evaluation in the Experimental Schools Projects

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Barbara Gross
Davis

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

April, 1976

Oren Glick, Ph.D.
Director, Program Evaluation
Franklin Pierce School District
Tacoma, Washington 98444

Table of Contents

	<u>Page</u>
I. Introduction to the Franklin Pierce Experimental Schools Project	1-4
II. The Evaluation Plan	4-6
III. Implementation of the plan	6-13
IV. The Range and Variety of Evaluation Activities	13-19
V. Some of the "Better" Activities and Products of Evaluation .	19-36
VI. Appendices	

Paper Prepared for Distribution in Conjunction
With the 1976 AERA Symposium Presentation: Internal
Evaluation in the Experimental Schools Projects
Oren Glick, Ph.D.
Director, Program Evaluation
Franklin Pierce School District

I. Introduction to the Franklin Pierce School District and its Experimental Schools Project.

The Franklin Pierce School District was one of the three initially funded Experimental Schools Projects. The project began operation in the summer of 1971 and is thus in its final phases and will be concluding as of June 30 of this year (1976).

The Franklin Pierce School District is located just south of Tacoma, Washington in a semi-urban, semi-rural area of some 19 square miles. The district has nine elementary schools, two junior and two senior high schools, with a total enrollment of about 7,800 students. The community is relatively placid with the schools, churches, and business clubs constituting the major forms of social organization. The district has enjoyed the relatively uncritical support of the community for the past five years and more. The district has a good history of passing special local levies which provide a major portion of the support for schools in the State of Washington. Indeed, in the face of increasing numbers of levy failures in the state in the past two or three years, the passage of the levy for the coming year's operation with a 73.1% majority last February must be considered remarkable. The only organized opposition to the District's programs over the past five years occurred this past winter and spring when a very small group of dissidents, opposed to federal support for education and claiming dissatisfaction with the schools, attempted to develop community support and exert pressure to eliminate specific portions

of the curriculum (e.g., MACOS and a secondary World History course). There is no evidence that they have met with much success and the large majority support of the special levy in February suggests that, if anything, their efforts boomeranged.

It is difficult to characterize the nature of the ESP in Franklin Pierce in a few words since it is so many different things. The typical rhetoric of the project proposal and other communications disseminated to the public present the central theme as "providing an appropriate environment for optimal learning for each student", with a special emphasis placed upon the "transient" student, the "high risk" student, and the "gifted" student. The emphasis of the project is not on testing novel developments, but rather on bringing together a variety of recently developed but "proven" educational models, programs, materials, equipment, methods, etc. and implementing them in an integrated and coherent fashion. One can best provide an overview of the nature of the project by listing the elements that have been brought into the district and/or supported through ESP:

A. Various forms and models of "individualized" instruction and management systems

1. IPI Math and Reading
2. Distar Math and Reading
3. Pullman P.E.
4. MACOS (Man a Course of Study)
5. People and Technology
6. Occupational Versatility
7. Programmed Maps and Study Skills
8. Systems 80
9. Teacher Developed Materials (LAPs)

10. Ginn 360
 11. BRL Sullivan
 12. Holt Rinehart and Winston Text and DataBank
 13. Criterion Reading
 14. Wisconsin Design (Word Attack and Study Skills)
 15. FPMSS (Franklin Pierce Mathematics Management Support System)
 16. Precision Math
 17. HEP (Hawaiian English Program)
 18. CEMREL (Central Midwestern Regional Educational Laboratory)
 19. SRA (Science Research Association)
 20. MAPs (Math Activity Packages)
 21. ESS (Elementary Science Study)
- B. Extended school year and interims
 - C. Career and Vocational Education
 - D. Work Experience
 - E. Personal Development, leisure time, and lifetime sports curricula
 - F. "Fifth-day" activities
 - G. Variable time in school
 - H. Instructional Aides
 - I. Adult Volunteers
 - J. Peer and Cross-age Tutoring
 - K. Community participation in educational planning and decision making
 - L. Environmental education
 - M. Drug education
 - N. International Education
 - O. A "balance" of "reactive", "active", and "interactive" instructional modes

- P. Teacher and student developed curricula
- Q. Year-round utilization of school buildings
- R. Parent Advisory Councils
- S. Cost-Effectiveness Model (Program Planning Budget Evaluation System-PPBES)
- T. Dissemination
- U. Improved Staffing Models
- V. New organizational structures
 - 1. Program Cabinet administration
 - 2. IGE schools
 - 3. "Open classroom" variations
- W. Teacher training
- X. Diagnostic-Placement testing in individualized programs with students progressing "at their own rates".

The implementation of this vast array of elements was intended to converge in its effects to produce the following kinds of outcomes for students:

- 1. Improved student attitudes toward school and the content of learning.
- 2. Improved student self-concepts of ability.
- 3. Variable and reduced time spent in basic skill learnings without reduced learning.
- 4. Reduction in the number of "high risk" students.
- 5. Reduction in discipline problems.

II. The Evaluation Plan

Prior to the advent of the ESP no institutionalized evaluation capability existed in the district. There was one relevant central office administrator position but it was unfilled for a year before the project began. University consultants were occasionally engaged

to assist in the evaluation of categorically funded projects. Achievement testing (ITES) occurred at certain grade levels and aptitude and interest testing was done in the secondary schools. However, the results of such testing were used in guidance and counseling functions rather than for program evaluation.

The evaluation plan identified three major goals:

1. To provide the Superintendent and his board with accurate information as to the progress and effectiveness of the programs undertaken by the project.
2. To set up and operate a program monitoring process that will provide teaching staff and district administrators with information needed in decision making.
3. To accumulate and periodically report to federal sources appropriate data relating to the progress of the program toward its objectives.

The major sets of activities and procedures proposed for pursuing these goals are summarized below:

1. District-wide (grades 2-11) achievement testing with the Comprehensive Test of Basic Skills (CTBS).
2. Every other year aptitude testing with the Short Form Test of Academic Aptitudes (SFTAA).
3. Development of procedures for tracking the progress of the district in implementing its programmatic objectives.
4. Assessment of staff attitudes, perceptions, and concerns regarding project implementation.
5. Provide technical assistance in curriculum development and selection so as to satisfy requirements of program evaluation.

6. Assist in the development and implementation of a program planning budget evaluation system (PPBES), including a cost-effectiveness model.
7. Develop and implement systems for monitoring individual student progress in specific curricula.
8. Develop and implement general criterion referenced diagnostic placement instruments for assessing and placing in-coming students and for reporting progress and placement of exiting students.
9. Develop and implement a district-wide student affective assessment program.
10. Provide an "audit function" on all programs funded under ESP.

III. Implementation of the Plan

A. Staffing and Budget

A considerable amount of negotiating occurred with regard to appropriate staffing and budget within the district and between the district and OE/NIE in the initial planning phase and again in arriving at the final 30-month contract. The results of these negotiations provided the following staffing over most of the five year period: A professional staff consisting of a director and three program evaluators; a support staff of three secretary-clerical persons, one data processing specialist, one-two computer programmers, and one keypunch operator. The total Evaluation budget has averaged about \$175,000.00 a year over the five-year project.

B. Constraints and Barriers in Implementing the Plan.

1. Establishing Adequate Data Processing Capabilities. A critical

element in the development of evaluation's role and function in the district was the establishment of an adequate data processing capability. As will be noted below, data of many kinds were generated at a tremendous rate. The Evaluation effort was very early finding it difficult to adequately organize, analyze, display, and interpret information to district personnel on a "timely" basis. A considerable amount of time and energy was devoted to this problem but, partly for reasons noted below, it was never satisfactorily resolved. The Evaluation Unit went through several phases of utilizing different equipment in an effort to adequately meet the data processing needs.

The Unit began its effort with desk calculators, including a programmable 1016 Marchant. During the first year a trial effort was made in contracting out certain data processing and analysis work to an adjacent school district that had a computer. Problems of access, control, and turnaround time made this alternative short lived. Unit record equipment (keypunches, card sorter, printer, collater, and calculator) were acquired in the second year of the project. These were a great asset but lacked the power and flexibility required for basic data reduction and for statistical analyses. In another effort to meet the data processing needs, a computer terminal was installed during the second project year. This alternative was not continued as it proved to be too costly and inadequate for handling large amounts of data. About the middle of the third project year an NCK Century 50 computer was obtained. An OpScan reader and attached keypunch were acquired at the same time.

While this equipment proved very useful in many respects, as will be noted below, the computer also brought on many problems and did not satisfactorily resolve the basic issue. It also gave the district a more expensive item of equipment than it believed it could sustain after the project. Consequently, in the last year of the project a transition was made to still another set of equipment, viz., the 9830A Hewlett-Packard Programmable Calculator with attached card reader and printer.

While the problems of establishing an adequate data processing capability had been a tremendous source of frustration to the Evaluation staff as well as to other district personnel it is at least gratifying to know that the project will conclude with the district having acquired a set of equipment that we believe is flexible and powerful and very appropriate for a district of the size of Franklin Pierce and consequently will meet a wide range of the district's data processing needs. This equipment consists of:

- a) the OpScan reader and attached keypunch, including a developed set of highly flexible response forms that can be used for gathering a very wide range of data from students, district personnel, and the community
- b) the Unit Record Equipment for organizing and sorting cards and obtaining listings, and
- c) the Hewlett-Packard equipment that will provide for adequate data reductions and summaries as well as a wide range of statistical analyses.

2. The multiplicity of goals and objectives, their frequent

vagueness and ambiguity.

Some sixty often overlapping project goals and objectives statements were contained in the funded proposal. They ranged from highly specific program implementation statements to remote student outcomes. For some of the objectives feasible data gathering procedures were successfully developed and put into operation. Others were developed but never got off the ground because they were too cumbersome and demanding of school personnel and were consequently resisted. Still other objectives successfully defied repeated efforts at meaningful measurement and were, for evaluation purposes, eventually abandoned.

3. Lack of clarity and consensus regarding the respective roles of and relationships between the Level I and Level II Evaluation teams.

Each of the Evaluation proposals submitted by Level I and Level II were approved and funded by OE. But the two proposals overlapped considerably in proposed evaluation functions and activities. Thus two independent agencies were committed to perform overlapping activities in the same time and place. It might be thought that the two parties under such conditions could resolve the dilemma themselves, that the total "pie" could be divided between them so that relationships and activities would be complementary, collaborative, and interdependent, or that the rhetoric of "formative" and "summative" evaluation would delineate separate but equal (while complementary) roles and functions.

It was argued by some of us at the time that this was not a reasonable expectation, that the assertion that "men of good will will work it out" was irrelevant. It wasn't a question of good or ill will. The problem was, rather, that the situation was structured so as to insure conflict and controversy. No very adequate resolution to this dilemma was ever achieved. There were isolated occasions of profitable interaction and collaboration and a reasonably workable agreement on information sharing and exchange was achieved. But generally the solution was for each party to go its separate and independent way and thus for the right and left hands not to know what each other was doing.

4. The problem of the art and science of evaluation research being yet in its infancy, a subject of controversy among the experts, and an unknown for the educational practitioner. This problem relates primarily to the issue of time; time to listen and communicate with the practitioners who are to be the beneficiaries of the evaluation activities; time to think and plan evaluation procedures and strategies; time to put them to an initial test before full scale implementation; time to plan and perform data reductions and analyses; time to plan how best to report the results and then to report them to the practitioners; time to pursue important but unforeseen methodological and substantive questions that arise in the course of the main activities. In all this, one is frequently reminded of the first (or is it second?) law of research formulated, tongue-in-cheek, by Arthur Bachrach: "Things take longer than they do."

5. Unlimited Expectations.

The third year of the project brought several significant changes for the evaluation unit:

- a. The director resigned to take a position with NIE.
- b. The NCR Century 50 computer was acquired.
- c. A new Director arrived on the scene.

The acquisition of the computer was as much bane as blessing. A number of unrealistic expectations, both within the district and within NIE, were established in the course of justifying the need for the computer and in obtaining approval for its acquisition. In addition to the data analysis requirements which it was initially intended to meet, the Evaluation Unit was now expected to

- a. Perform the scheduling, grade reporting and attendance record keeping for the secondary schools.
- b. Prepare to take on the business office functions (accounts payable and receivable and payroll).
- c. Develop and implement a comprehensive student information system.
- d. Score and report standardized achievement and aptitude testing.
- e. Develop and implement student progress tracking systems for individual curricula.
- f. Score and report criterion referenced tests developed for assessing achievement of locally determined goals and objectives.
- g. Perform services for out-of-district clients so that by the end of the project the unit would be approximately 50% self-supporting.

While there seemed to be some improvement in the presentation

and interpretation of data by the third year of the project, the problem of "timeliness" of reporting only worsened. Over criticism of evaluation appeared to increase throughout the district. Morale within the Evaluation Unit became a serious problem as staff felt overwhelmed by the scope of the activities being undertaken and repeatedly frustrated in their efforts to provide satisfactory services to the schools. A management information system was implemented, numerous activities were curtailed, and priorities were set and reviewed on a recurring basis. But it was a case of too little too late.

In the fourth year of the project an Evaluation Committee was formed, made up of several principals, teachers, and counselors, and chaired by the Level I Director. This committee was to review the activities of Evaluation, serve as an advisory group, and make recommendations for Evaluation activities for the final year of the project. The committee conducted a survey in which the recipients of particular evaluation services throughout the district judged the usefulness of those services to the functioning of their programs and indicated their attitude toward the district financially supporting the continuation of those services after the project was over. Attachment A of the Committee's report listed 20 major sets of activities that the Evaluation Unit was either operating, had under development, or had under consideration for development. That listing is reproduced here since it provides a good summary of Evaluation's major activities at that point in time, i.e., as of the middle of the fourth year.

Since disillusionment with the computer was so wide-spread, a management consultant firm, with special expertise in the analysis of computer operations, was contracted to make an assessment of the

utilization of the computer and to make recommendations to the district regarding possible changes in its utilization for the duration of the project and for its post-project utilization in the district. To make a long story short, the consulting firm recommended that the district not retain the computer. It was then decided not to resume, on the computer, the operations of any of the information systems when school resumed in the fall of 1975. Those operations would either be dropped at the end of the 74-75 school year or transferred to other systems. The computer would be retained only long enough to get all of the data files in order and perform the necessary analyses and reports for evaluation requirements.

Consequently, still another phase was entered in the district's effort to obtain a satisfactory data processing capability. The district decided to retain the unit record equipment, to purchase the OpScan reader and attached keypunch, and moved in the winter of 1975-76 to purchase a Hewlett-Packard 9830 Programmable Calculator with attached card reader and printer.

IV. The Range and Variety of Evaluation Activities.

In an effort to be responsive to the varied information needs of the district, as well as to fulfill the initially projected evaluation plan, a tremendous amount of data gathering activity was initiated during the first 30 months of the project. The Evaluation report for the first 30 month period attempted to summarize and review the effort to date. Data that had been gathered fell into two general sets: 1) Those that were general and district-wide, not linked specifically to any particular project component, event, or activity. Included here, for example, were district-wide assessments of student achievement, attitudes, and self-concepts of ability. 2) Those that

were event or activity specific and which called for sampling of particular population subsets involved in those specific events or activities. The first 30 month report summarized why, how, and when each kind of data was collected, and by whom and from whom it was collected. Each identified data set was also classified into one of four categories indicating the status of its disposition: reported, processed, partially processed, or stored. The data gathered to that point in time was as follows:

A. General, District-wide Data

1. California Test of Basic Skills (three administrations)
2. Short Form Test of Academic Aptitude (two administrations)
3. Student Attitudes Toward School (four administrations)
 - a. Grades 1-3: "Feelings About School"
 - b. Grades 4-12: "Pupil Opinion Questionnaire" for general attitudes
 - c. Grades 4-12: "Semantic Differential Scales for Attitudes Toward Specific Subject Areas"
4. Student Self-Concept of Ability (four administrations)
 - a. Grades 1-6: "About Myself"
 - b. Grades 7-12: "Self-concept of Ability: General"
 - c. Grades 7-12: "Self-concept of Ability: Specific Subjects"
5. Staff Attitudes Toward ESP (six administrations)
6. Community Survey (one administration)

(Level I collaborated with Level II in this extensive effort involving a combination of questionnaire and interview)
7. Data for Cost-Effectiveness Analyses (two administrations)

(Number of teacher hours spent in each course; number of students in each course.)

B. Data Specific to Program Activities

1. Curriculum development (three times).

(Identification of newly developed and revised curriculum through interviews, inspection of contracts, and curriculum development products.)

2. Staff attitude toward the interims

(Questionnaire administration to the faculty of the two senior high schools and one of the junior high schools.)

3. Data on student completion of objectives in winter Interim courses for 1972 and 1973.

4. Combination of observations, checklists, and teacher interview to assess project objectives with respect to the winter Interims for 1972 and 1973.

5. Student attitudes toward winter Interims for 1972 and 1973 and the summer Interim, 1972 (in collaboration with Level II).

6. Questionnaire assessment of parent attitudes toward the winter Interims for 1972 and 1973.

7. Questionnaire assessments of the percent of time teachers engaged in each of three instructional modes (two administrations).

8. Identification of special instructional activities and programs for "high risk" students through inspection of school records and student interviews (three times).

9. Examination of school records to determine amount of extended school building use (other than regular day).

10. Post meeting teacher workshop data.

(Questionnaires were administered to obtain participant's feedback on the strengths and weaknesses of numerous training workshops.)

11. Examination of records and interviews with work experience counselors to determine the number of students in work experience programs (nine times).

12. An examination of records and teacher interviews to determine adult/student ratios for each course (five times).

13. Inspection of school records and interviews with teachers and principals to determine organization of fifth-day activities and course offerings (eight times).

14. Questionnaires administered to participants of community Charrette to assess community reactions to ESP components (twice, and in collaboration with Level II.)

15. Questionnaire administration to students regarding attitude toward various aspects of the Occupational Versatility program, assumption of responsibility by students and use of information resources by students.

A total of 26 written Evaluation reports were produced in the first 30 months. However, of the 74 data gathering episodes that were documented, 28 (38%) of them had not resulted in a report to anyone. And for the nearly two-thirds of data gathering operations that were reported, there were frequent problems. One major problem appeared to be "timeliness". Reports frequently were not available when decisions that could have used the reports were being made.

But a second problem had also become apparent: The reports were often of disappointing value in decision making even when they were, or if they had been, timely. Several factors could be identified that appeared to contribute to this problem:

1. In some cases the problem appeared to be due to the lack of cogency inherent in the data themselves relative to the questions or issues at hand, or to the decisions to be made. We simply hadn't generated very good data. Data gathering was often undertaken on short notice with a minimal amount of planning. The particular purposes to which the results would be applied were often not well thought through or articulated. Nor were the data gathering procedures always carefully constructed so as to insure relevance to a particular purpose or purposes.

It was unclear just how evaluation questions and issues were to be formulated and who was to be responsible for what those questions and issues were to be. It was easy for evaluation team members to say to administrators and teachers,

"You just tell us what you want to know or what information you need in order to make better decisions in operating your program and we'll help you get it. But we can't help you if you don't tell us what you need to know." The program people were typically not prepared to answer that question very clearly. And the well intended initiatives of evaluators in communications of that form did not encourage them to try, perhaps because they carried the implicit message that the practitioner should be able to answer the question. This set the stage for those who did try to probably not be very reinforced for doing so. Chances were they would discover (from the evaluators, of course) that their questions as stated couldn't be answered for one reason or another. The alternative of the Evaluator formulating the questions or issues, while having a higher probability of being answerable, also had a high probability of generating answers to questions few besides him/herself were interested in.

The third alternative, of course, is for questions and issues to be a product of collaboration between the practitioner and the evaluator. How situations and occasions for the necessary interaction to occur can best be arranged is problematic. But the respective roles are, I believe, relatively clear: the practitioner must express his/her ideas, concerns, observations, questions, understandings -- what have you -- that derive from his/her direct experience in the program. Such expressions are likely to be, in varying degrees, unclear, vague, ambiguous, and incomplete. That's all right.

We shouldn't expect it to be otherwise. It should be one of the skills of the evaluator to listen carefully and, in the course of continued interaction, arrive at a formulation of the issues and questions so that they are sufficiently clear and answerable. Furthermore, the final formulation should be more satisfying to the practitioner as an expression of his/her need than were their initial statements.

2. In some cases the problem of "relevance" appeared to be due to how the data were organized and displayed. Often they were not sufficiently or properly reduced and analyzed so as to bear on particular questions of interest. In these cases the decision maker was sometimes confronted with an overload of discreet information that was, perhaps, "interesting" but was not interpreted and not in such a form that the decision maker could readily draw the desired or necessary inferences. Within the evaluation team there were different points of view and sometimes conflict over whether we should simply "get the information out" as soon as possible in most any form or whether we should take the time to reduce, organize, analyze, and display the information in a way that would clearly reflect on particular issues, questions, or decisions.
3. A third aspect of the problem of "relevance" became apparent with the increasing clarity and persistence of the call for program specific evaluation. District personnel were asking questions, and decisions were going to be made, about particular curriculum programs. While it appeared obvious in retrospect that we should have anticipated this need, it was not sufficiently apparent, nor sufficiently provided for,

in the early planning and operational stages of the evaluation effort. We had gathered data on a massive scale from students for some time. These data were coded for school identification, grade level, and sex (usually) but not for program. Consequently, data that had been gathered district wide could not readily be assembled so as to provide for comparisons among programs.

V. Some of Level I Evaluation's "better" activities and products.

The items presented below reflect, of course, my own perceptions and values. Other people knowledgeable in this situation might have made other selections. I have, however, tried to select some of those activities or products which appeared to have some impact within the district, or for which definite feedback was received, or which appeared to be of some general interest and applicability.

A. Studies that attempted to show the impact, or relative effectiveness, of various project activities or components.

1. An assessment of the effects of the interim on student achievement in language arts and arithmetic (Evaluation Note #19).

One of the claims of the project was that basic skill achievement could be maintained without loss and in less time with improved programming. In that case it would be possible to introduce a three to four week "Interim" of predominantly "personal development" courses without incurring negative consequences in achievement. At the same time, of course, there would accrue the benefit of the personal development learning for students and, hopefully, enhanced student attitudes toward school. One of the two junior high schools initiated the interim three years after the other one. Consequently, in

the spring of 1974 there were two sets of ninth graders in the district, one of which had spent a total of 45 fewer days in basic skill instruction over the preceding three years than had the other. The two groups were compared by means of an analysis of covariance on the spring, 1974 language arts and arithmetic sub-scales of the CTBS using the 1971 scores as the covariate. On each of the eight adjusted measures, the school with reduced time in basic skill instruction was superior to the other school. Six of the eight comparisons were statistically significant. For proponents of the Interim concept this was, of course, welcome news. Maintaining, or even enhancing, basic skill achievement was apparently not necessarily incompatible with simultaneously doing other things.

- B. Does implementation of the Grace Nash Music program enhance student attitudes toward music?

This question was prompted by discussions with the District's Coordinator of Music, regarding the possible impact of the Grace Nash Music program. Data were available on student attitudes toward music prior to the initiation of the Grace Nash program and again two years after implementation. Measures were also available on the extent of implementation of the Grace Nash program in the various elementary schools. If implementation of the Grace Nash program has favorable effects on student attitudes toward music, then the correlation between the measures of implementation and the 1975 post test of student attitudes toward music should be positive. And they should be greater than the correlations with the 1973 pre-test measures which

should not differ significantly from zero. The results indeed corresponded to these expectations. The findings were thus consistent with the hypothesis that implementation of the Grace Nash program enhances student attitudes toward music: the more a school implemented the program, the more favorable were subsequent student attitudes toward music. But correlation per se, we know, does not mean causation. Perhaps the schools showing greater implementation were generally "on the move" more than other schools during the time period in question. Perhaps greater implementation of Grace Nash was only a part of the innovative developments occurring in all subject matter areas. If this is true we should find student attitudes toward other subject matter areas as well as toward music, showing positive correlations with the measure of implementation of the Grace Nash program. These correlations were calculated and found to be near zero. Thus the interpretation was supported that implementation of the Grace Nash program enhanced student attitudes toward music.

- C. An analysis by grade level of the elementary school CTBS scores relative to the ability levels of the schools.

A sore point among some of the schools in our portrayal of the CTBS data was that it failed to take into account the varying ability levels among the schools. While the California Test Bureau makes available the option of "anticipated achievement scores", their regression formula at that time was not being made available and we were doing our own scoring and reporting. However, we did have I.Q. data. Consequently,

our approach to this question was to regress individual school average grade level CTBS scores on average grade level ability (IQ) scores thus generating a predicted average grade level CTBS score on each measure at each grade level for each school. For each grade level and subscale, each school could then be plotted showing its deviation from the regression line. It was immediately apparent from inspecting such a graph which schools are performing better or worse than expected.

- D. An analysis of secondary school student attitudes toward their courses -- Spring, 1975.

During the first several years of the project student attitudes toward courses and self-concepts of ability in subject matter areas were measured by Likert type attitude scales and semantic differential techniques. Thus relatively large numbers of items were involved which made the administration of the instruments obtrusive and cumbersome and generated a tremendous volume of data, more than could be quickly and conveniently processed and reported back. A very simple approach to the measurement of student attitudes was experimented with in the Interim and in the fourth year of the project these developments replaced the more cumbersome methods that were previously used in the regular term. The instrument is reproduced as Appendix B. Students respond on an OpScan answer form which is automatically transformed to punched cards and then readily processed. The instrument is very easily and quickly administered by teachers. Indeed, while teachers previously complained

about the length of the affective instruments, some teachers now complained that students answered these items so quickly they they couldn't possibly have any value. Nevertheless, an analysis of variance showed that there were highly significant differences among the six curriculum career streams, that these differences varied among schools, that there were significant differences among courses, among the three attitude questions, and that the differences among the questions varied among the six streams. The significant interaction between schools and career streams, for example, is portrayed in Appendix C.

A simple data organization technique, which we have found can frequently be used to great advantage since it greatly facilitates a "meaningful" reading of the data, is to show the entities being measured in rank order. This is particularly true when many entities are involved. In Appendix C1 we have reproduced the first page of the 11 pages of student attitudes toward the 266 courses offered in the secondary schools. The courses are listed on the basis of the ranking of the average rating over all three attitude questions as reflected in the right hand column. The mean rating on each question, and the rank of that rating, is also portrayed. Such a ranking of the courses shows very readily which courses are associated with the most favorable and the most unfavorable student attitudes.

E. The Comprehensive Survival Skills Test in Math.

The senior high schools developed a math program designed to ensure that every graduating senior possessed the minimal math skills required for "successful living" in our society.

about the length of the affective instruments, some teachers now complained that students answered these items so quickly they they couldn't possibly have any value. Nevertheless, an analysis of variance showed that there were highly significant differences among the six curriculum career streams, that these differences varied among schools, that there were significant differences among courses, among the three attitude questions, and that the differences among the questions varied among the six streams. The significant interaction between schools and career streams, for example, is portrayed in Appendix C.

A simple data organization technique, which we have found can frequently be used to great advantage since it greatly facilitates a "meaningful" reading of the data, is to show the entities being measured in rank order. This is particularly true when many entities are involved. In Appendix C1 we have reproduced the first page of the 11 pages of student attitudes toward the 266 courses offered in the secondary schools. The courses are listed on the basis of the ranking of the average rating over all three attitude questions as reflected in the right hand column. The mean rating on each question, and the rank of that rating, is also portrayed. Such a ranking of the courses shows very readily which courses are associated with the most favorable and the most unfavorable student attitudes.

E. The Comprehensive Survival Skills Test in Math.

The senior high schools developed a math program designed to ensure that every graduating senior possessed the minimal math skills required for "successful living" in our society.

The program called for periodic testing with test forms constructed by the math departments. The evaluation unit provides several services in conjunction with this program: 1) test scoring, 2) printing of student report cards with subscale scores and a message of congratulations or regrets with further instructions for what the student should do next, 3) printing of test result stickers for attaching to each students' file, 4) a response analysis report showing for each question, the number and percentage of students selecting each alternative, and 5) a student response report showing, for each student, the items responded to correctly and incorrectly. These services will be continued in the post-ESP era and are being expanded into the lower grade levels and into subject areas other than math. The Evaluation Unit also performed analyses showing that the total score on the CSST correlated .85 with the total arithmetic score on the CTBS. Based upon this finding a regression analysis was performed, regressing CTBS on the CSST. Predicted CTBS scores (converted to grade equivalents) were reported for the range of CSST scores which enhanced the framework for understanding the CSST scores. The passing criterion of 75% correct on the CSST was found to fall right at the 12.0 CTBS grade equivalent score, a finding which stimulated considerable debate about what constituted a reasonable or desirable criterion for "passing".

F. A Cost-Effectiveness Model.

We evolved the concept of a cost-effective matrix (See Appendix D) which has had a great deal of heuristic value

but which, operationally, is still far from satisfactory.

I will discuss a number of aspects of this development.

The implementation of a cost-effectiveness model required some system for tracking cost factors for individual programs. General costs of personnel were readily available from the business office. Information on the distribution of each teacher's time over the programs was more difficult to come by. We settled for teacher estimates obtained by means of a questionnaire.

A system was also developed for tracking nonsalary costs. A taxonomy of courses was developed providing six categories (social studies, math, science, communications, personal development, and career education) called "career streams" within which every course was located. Every course was assigned a four-digit number for tracking purposes. Every purchase order issued by a budget administrator had to carry the four-digit course number or the business office wouldn't process it. While this development had considerable merit it also carried several serious problems:

- a. It was generally difficult for budget administrators to allocate their purchases to specific courses. Many purchases were of necessity distributed over a range of courses. Four-digit course numbers were therefore supplied to cover purchases that were general within streams and for purchases that were general across the total curriculum. It was, of course, easier to use the more general numbers. Budget administrators varied in the extent to which they used the course

specific ones. For the 1974-75 year only about 25% of the purchase orders carried course specific numbers, leaving 75% of the purchases being distributed over a total stream or over all courses. When you consider the fact that non-salary costs of education are about 15% of the total costs and that our cost-tracking system captured only 25% of that, it is apparent that a fairly elaborate system had been put into operation that accounted for only 4% of total expenditures.

- b. A second difficulty arises in connection with pinpointing the time period within which the benefits of a given expenditure are to be realized. The "effectiveness" part of the model measures student growth or outcomes within a designated time frame, usually one calendar year. Costs assigned to that growth should be those incurred during that same time frame. More specifically, the costs of the equipment, supplies, and materials actually utilized within that time frame are the ones to be tracked. Costs of equipment can be averaged or pro-rated by some formula over the estimated lifetime of the equipment. But what about supplies and materials? The purchase orders carry an issue date. What can that tell us about the time frame within which the products of the purchase will be utilized? For example, should purchase orders carrying a May date be assigned to that same year or to the succeeding year? Or should it be split in some proportion between the two years? This

would seem to be a general problem for any cost-effectiveness model: How in an on-going district operation can you accurately assign program costs to the same time frame in which student growth is being assessed?

There is a further difficulty with the basic concept of cost-effectiveness systems. Presumably the purpose of designing and operating cost-effectiveness systems is to improve decision making at the school district level so as to arrive at the best educational programs for the least cost. It is generally assumed that the decision making necessary to arrive at that goal resides primarily with school administrators. On closer examination it becomes apparent that this may not be the case. Costs of education are primarily a function of three factors: teacher salary, teacher/student ratios, and time in program. Salaries and teacher/student ratios are negotiated. Novel and less expensive staffing arrangements are resisted by teacher organizations. Time in school is legislated. The whole issue of educational costs has thus become predominantly a political issue. Effective action and decision making reside largely outside the domain of operational management.

But there may be some silver linings to pursue in this cloudy picture. While the general parameters determining overall costs are set through various political processes, there is still room for considerable variability within (or sometimes around) them. For example: 1) Overall time in school may be legislated, but how student time is distributed

within school is not. Through ESP, the Franklin Pierce district has significantly varied the amount of time students spend in particular programs by implementing "fifth day" and "interim" concepts. 2) The Franklin Pierce teachers organization successfully blocked the inclusion of a differentiated staffing plan in the original ESP proposal but the district nevertheless had some schools adopting the IGE and other unit organizational plans and employed teaching interns and numerous instructional aides.

G. A General Model for Program Evaluation.

What does all this suggest in the way of program evaluation activities within a school district? Suppose we let the cost dimension of the cost-effective matrix become any variable (or combination of variables), including indicators of cost as suggested above, that might be considered to have an effect on student growth or outcome measures of interest. To the extent that variation in such variables of interest is found to occur naturally (or can be deliberately arranged) we have an opportunity to assess their relationship with (or even their effect upon) the student growth measures. The data analytic techniques best suited for this situation are those of general multiple linear regression.¹ This brings us to the present stage of our evolution in evaluation at the Franklin Pierce school district.

We are currently implementing a concept of program evaluation that will probably constitute the major legacy of ESP evaluation in the district. Its continued development and assessment in

the post-ESP era will be sustained through a Title IVC grant. The general approach is described in an Evaluation Technical Report titled, "A General Design for Evaluating the Relative Effectiveness of Basic Skills Programs within a School District." I will briefly describe the approach. A schematic that is useful in understanding the design may be found in Appendix E.

The design consists of two interrelated parts. Part I addresses the question, "What is the relative effectiveness of each of the instructional units operating in the district?" Part II addresses the question, "What factors might account for the differences in relative effectiveness among the instructional units?"

Instructional activities occur in what may be called "instructional groups". That is, some number of students and instructional staff come together at specified times, in specified places, and engage in some kind of identifiable instructional activity. Such activities, involving predominantly the same students, have a certain duration in time (academic year, semester, interim month). They have a certain frequency of occurrence within that time period, and each occurrence has a typical length of instruction time. Instructional groups are thus real-life units that exist in space and time. They are the basic organizational blocks of which the total instructional program is built. It is those building blocks, and the characteristics associated with them, that determine the strengths and/or weaknesses of the entire program. Consequently, it is argued that any viable design for program evaluation must use

the instructional group as the basic unit of analysis.

Instructional groups are designated by capital letters in the figure in Appendix E and appear in next to the last column.

The level of performance characteristic of any instructional group may be represented by the average measure of the members of that group. Such a mean score obtained near the close of an academic instructional period represents the level of growth for that group at that point in time. These values are represented in the far right hand column of the figure in Appendix E.

Taken by themselves, however, these mean scores do not accurately reflect for each group the amount of growth or change associated with the just concluding instructional activity. Such scores reflect in addition, e.g., the level of growth of each group before the particular phase of instruction involved was initiated. It thus becomes necessary to acquire a pre-instruction mean score for each instructional group. In a district with an annual testing program this can be done by reconstituting each of the groups in the previous year's testing and calculating the mean score for that point in time. Such a procedure is reflected in the figure in Appendix E. The column labeled "Students" identifies the students of each instructional group with brackets pointing to both a pre and post test average score.

The necessary data for Part I of the design have now been identified. The statistical analysis to be performed on the pre and post test means is a linear regression analysis. This

analysis removes from the post test mean scores that variation which was already present in the set of pre-test mean scores. In effect, it sets all of the instructional groups at a starting point of zero. After this is done, the remaining variance in the post test means more closely reflects that which is associated with characteristics of the just concluding instruction. The rank ordering of the instructional units on the basis of the adjusted scores shows the relative effectiveness of those units for the instructional period under consideration with respect to the measure of interest.

It may be helpful to digress for a moment and discuss the nature of the general analysis problem. Just as individual test scores differ, reflecting variations in levels of individual skill acquisition, so, too, mean scores of the instructional groups will differ, reflecting different levels of skill acquisition for each group. What we are after is an explanation of that variation. Why are some group means relatively high, others relatively low? It has already been noted that part of the variation reflects that which was already there before the instruction in question began. We can remove that source of variation, as indicated above. This will leave a certain proportion of the original variation yet to be accounted for.

It may be observed that the remaining columns in the figure in Appendix E are identified as "Potential Input Variables". This set can include any identifiable characteristics associated with the instructional groups. While the variables included

are considered to be likely candidates, the set is not intended to be exhaustive. Several of these variables, or variable sets, will now be discussed.

In order to draw conclusions regarding the relative effectiveness of programs, each instructional group must be identified in terms of the program involved. Program effectiveness comparisons can be made among any programs that are so identified. All that is required is that each instructional group be assigned to one and only one of a set of program categories.

The regression analysis will show what additional portion of the total variance in the set of post test means is associated with the various programs that have been identified. It will tell if that amount of variation is statistically significant. It is important to understand that the variation associated with programs is determined after the variance attributable to pretest differences has been removed. The different programs can then be rank ordered from high to low to see which ones are relatively more or less effective.

Any reliable information on the characteristics of instruction within the instructional groups can be examined for its possible contribution to the variance in the set of post test means. If multiple instructional characteristics have been identified, the proportion of variance due to the whole set, as well as to each characteristic individually, can be determined. The minimal requirement, again, is that the information regarding each instructional characteristic

to be included in the analysis be available for each of the instructional groups.² The information contained in the Instructional Characteristics Profile used in the Franklin Pierce School District could be used here.

It would be redundant to continue discussing each of the potential input variables shown in the figure of Appendix E. The general model should be clear: Any information linked to the instructional groups can potentially be included in the analysis. The results provide a test of whether or not any given factor is associated significantly with the outcome measure of interest.

The design presented here provides two important sources of information for program planning and change. One comes from Part I of the analysis in which the most and least effective instructional units have been identified. A comparison of these sets should suggest certain corrective actions that can be taken to improve program effectiveness. The second source comes from factors in the input variable set that have been found to contribute significantly to the outcome variable. To the extent that significant input variables are manipulable, instructional changes can be made that have a high probability of increasing student achievement. In both cases, the effectiveness of any instituted changes is put to an empirical test in the year in which they are implemented. In this way an ongoing annual cycle can be established: Each year a search is made for those factors that contribute significantly to explaining the outcome measure(s) of interest. The findings

provide input into planning for any changes to be instituted the subsequent year. And each year the effectiveness of changes made on the basis of the previous years results is determined.

The total design thus establishes an iterative process that will move a district toward maximal effectiveness and efficiency in achieving its objectives:

- Year 1: 1) instructional activities are identified and ordered in terms of their relative effectiveness.
- 2) factors or characteristics of instruction associated with those differences, and that might account for those differences, are identified.
- Year 2: 1) corrective actions, based upon the first year's results, are taken as needed.
- 2) steps (1) and (2) of Year 1 are repeated, but they now incorporate an empirical test of the effectiveness of the changes instituted in Year 2.
- Year 3: The cycle is now established, as Year 3 replaces Year 2.

H. The Instructional Characteristics Profile.

A frequent problem in evaluation research is that of adequately describing the nature of the programs, or differences among the programs, that are being evaluated. While certain characteristics of different programs may be generally known, frequently no data are available to determine what particular aspects of program differences might be responsible for varying

student outcomes when such are discovered. We have taken a number of approaches to this problem ranging from open ended descriptions by teachers to the structured check list of 33 instructional characteristics shown in Appendix F. We have had teachers merely check all those characteristics that are "typical" of their instruction for each course they teach. We are now having them check one of five alternatives indicating the percentage of instructional sessions in which each characteristic is present.

Each of the 33 characteristics has been classified in terms of whether the ESP goals and objectives called for it to "increase" or "decrease". All district courses have been classified as "C" (continued unchanged from before ESP) or "R" (revised or added through ESP). Analyses of these data have shown that for most curricular areas many of the instructional characteristics are responded to differently by teachers of C and R courses, and nearly always in the direction predicted from the project goals and objectives.

The instructional characteristics profile data will also be a source of independent variables as we pursue the general evaluation design described above.

I. Source Book on Courses.

A major product of the ESP evaluation effort will be a "Source Book on Courses". This will be assembled in loose-leaf form for periodic updating. It will consist of two pages for each course, K-12, offered in the district. The first page consists of descriptive information on the course. A

completed sample page may be found in Appendix F. The second page will report student outcomes, both achievement and affective, for the course, and certain cost factors. The format for that page has not been finalized as yet. The student outcome data will be that derived from Part I of the general evaluation design described above.

Footnotes

¹Two excellent sources for the application of multiple linear regression technique to data analysis problems encountered in program evaluation are Applied Multiple Regression/Correlation Analyses for the Behavioral Sciences (1975) by Jacob and Patricia Cohen and Multiple Regression in Behavioral Research (1973) by Fred Kerlinger and E. Pedhazur.

²It actually is not a requirement for the analysis that information on any particular variable be available for each unit of analysis. See Cohen and Cohen, Chapter 7 for procedures on how to include the presence or absence of data as an independent variable in the analysis.

Appendix A

List of Identified Evaluation Activities

I. Set I: Activities now operational or in the process of being implemented.

101 STANDARDIZED ACHIEVEMENT TESTING (CTBS)

1. Ordering necessary Examiners Manuals, test booklets, and answer sheets, preparation of test packets for administration.
2. Test scoring and reporting (Administrative Summaries; Class Record Reports; Individual Student Records).
3. Special analyses and reports on the CTBS data (Longitudinal analyses; rank order listings of students within grade levels on Reading, Language, and Arithmetic; listings for each student of band designations (1 to 6 representing grade levels above or below grade level for successive administrations of the test); comparisons of selected programs, etc. Reports on "high risk" students; reports on transient students.

102 BIENNIAL APTITUDE OR INTELLIGENCE TESTING

1. Ordering Examiners Manuals, Test Booklets, and Answer Sheets; preparation of test packets for administration.
2. Test scoring and reporting. Verbal, Nonverbal, and Total IQ reported for each student; individual student raw scores, verbal, nonverbal, and total IQ scores listed by grade level within each school; means and standard deviations by grade level.

103 STUDENT PROGRESS TRACKING SYSTEMS: (IPI AND MBO AT FORD)

1. Weekly reporting of cumulative individual student progress through specified objectives of program(s).

104 STUDENT INFORMATION SYSTEM: (SIS)

1. Printing of School Directories, updated monthly, or bimonthly (showing student ID number, name, parent's name, address, and phone number; alphabetized within grade level).
2. Printing of 5 X 8 information cards for elementary schools (New at start of each year; monthly or bimonthly updates).
3. Printing of mailing address labels (district-wide, by school, or grade within school).
4. Enrollment recording and reporting (First of year and monthly or bimonthly updating on new enrollees, withdrawals, total enrollment counts by grade within school.)
5. Reports or summaries of information contained on the enrollment form available on request.

105 COMPREHENSIVE SURVIVAL SKILLS TEST (CSST)

1. Answer sheet preparation for administration of the test; scoring of the test; report card printouts; printing of address labels for mailing report cards; listing of test results by student; response analyses by item.

2. Analyses of test data and written reports.

106 INTERIM EVALUATIONS

1. Report showing a) student attitudes about the Interim and the courses they take, b) students' perception of parental feelings about the interim, c) display of objectives for each interim course and data on percent of students achieving objectives.

107 PARTICIPATION IN CATEGORICALLY FUNDED PROGRAMS: TITLE I, TITLE III

1. Needs assessment, formulation of objectives, and development of the evaluation design required for the application for such funds.
2. Provide listings of students' CTBS scores for the identification of eligible Title I and gifted students. Conduct the program evaluations: monitor data gathering; analyze the data and write reports.
3. Monitor program compliance with federal and state guidelines and regulations.
4. Participation in state Title I and Title III meetings.

108 BUDGET PREPARATION AND REPORTS

1. Budgetary information listings for preparation of preliminary and final budgets; printing of budgets.
2. Monthly budget listings by location with available balance column corrected for "transfers".

109 EVALUATION OF TEACHER TRAINING AND CURRICULUM DEVELOPMENT

1. Attendance at workshop sessions by Evaluation staff member.
2. Construction of feedback instruments to which workshop participants respond (may be pre-workshop, in-progress, post-workshop, follow-up) and written reports based on feedback data.

110 MUSIC INVENTORY

1. Annual printing of musical instrument and equipment inventory.

111 TRANSPORTATION DATA REPORTS

1. Periodic summaries of data on district transportation: costs, miles, programs, destinations, etc.

II. Set II: Activities under development

212 TESTING FOR LOCALLY DETERMINED GOALS AND OBJECTIVES

1. Assist in the construction, selection, and evaluation of proposed test items for testing locally determined objectives.
2. Assist in assembling sets of tests (pretests, posttests, alternate forms, etc.) for assessing achievement of locally determined objectives.
3. Score locally developed tests and report individual student results and group summaries.

4. Prepare reports on analyses of test results including interpretations regarding program effectiveness.
- 213 COST-EFFECTIVENESS ANALYSES OF PROGRAM ACTIVITIES
1. Calculate and print for each school an annual listing of cost-per-student-instructional-hour for every program activity (course).
 2. For all Reading and Math program activities calculate and print annually measures of student growth for each program activity using relevant CTBS scores, attitude scores, and self-concept-of-ability scores.
 3. Assist Teachers, Principals, and Central Office Administrators in the interpretation of the cost and effectiveness data for purposes of decisions regarding the continuation, modification, or termination of particular program activities.
- 214 PROGRAM ACTIVITY DESCRIPTIONS
1. Annually print for each school an updated "Course Characteristics Profile Report", showing the profiles of every program activity (course) as described by the teachers teaching them. (This is a check list of instructional characteristics that teachers would check as being true for each course they teach.)
- 215 STUDENT ATTITUDES TOWARD COURSES AND SELF-CONCEPT OF ABILITY (TO SUCCEED) IN COURSES
1. Using data gathering instruments of the kind used in this year's interim, gather data and issue reports each term summarizing student attitudes toward courses, and judgments of their ability to succeed in courses.
- 216 CATALOGUE OF FRANKLIN PIERCE PROGRAM ACTIVITIES
1. Prepare annually a Franklin Pierce School District Program catalogue that would pull together and summarize, for each program activity (course) in the district, the "Course Characteristics Profile" information and the cost-effectiveness information.
- 217 STUDENT PROGRESS TRACKING SYSTEM (HEP)
1. Reports: Detailed individual student profiles showing dates program objectives were begun and completed; class summaries; summaries for objectives.

STUDENT QUESTIONNAIRE: COURSE ATTITUDES

Students: Please do not write on this sheet. Responses are to be marked only on the answer sheet provided.

1. I would rate my effort in this course as:

Very Low 1 2 3 4 5 Very High

2. I feel that the information learned in this course is:

Not at all
Useful 1 2 3 4 5 Very Useful

3. On most days this course is:

Very Very
Boring 1 2 3 4 5 Interesting

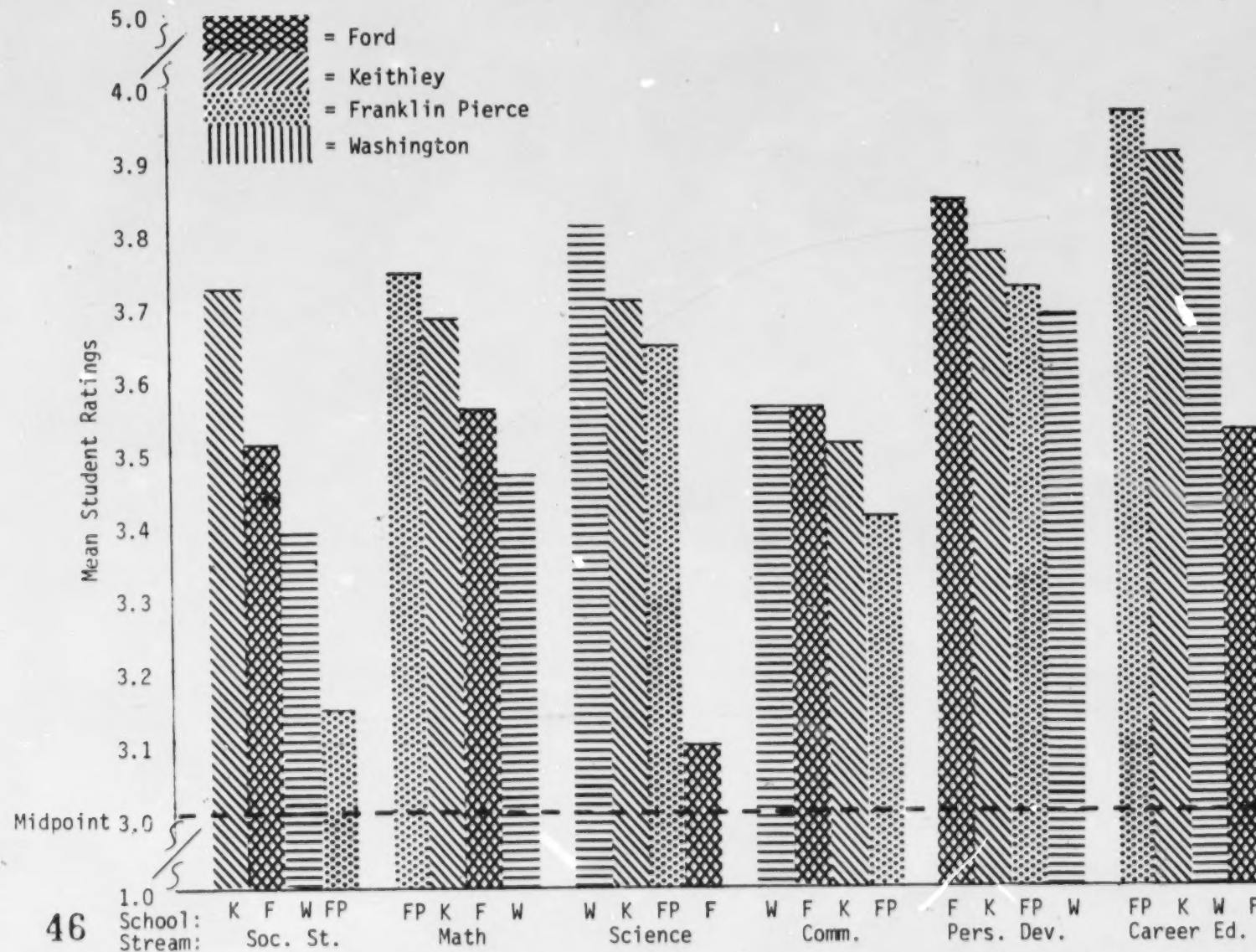
4. Compared with other students in this course, how do you rate your ability in the subject matter of this course? I am,

1. among the poorest.
2. below average.
3. average.
4. above average.
5. among the best.

5. Compared to other courses you have taken or are now taking in junior or senior high school, how do you rate your ability in the subject matter of this course? I am,

1. poorer than in any other course.
2. not as good as in most other courses.
3. about the same as in most other courses.
4. better than in most other courses.
5. better than in any other course.

Appendix C
Differences Among Schools in Student Attitudes Toward Courses Within Each
Career Stream
($p < .001$)



Appendix C1

Secondary School Student Attitudes Toward 1974 - 1975 Courses:
Mean Ratings and Rankings on the Three Attitudinal Components of
"Effort", "Usefulness of Information", and "Daily Interest Level",
Sequenced by Total Attitude Mean Ratings.

Course Title	Effort (E)		Useful (U)		Interest (I)		Total Mean $\frac{E+U+I}{3}$	
	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
23-6351 Occupational Exploration (Y)	4.500	10.0	4.667	19.0	4.667	1.0	4.611	1.0
26-4505 Advanced Public Speaking (S)	4.625	3.0	4.750	9.5	4.250	7.0	4.542	2.0
26-6593 Model Office Simulation (Y, 2 hrs.)	4.722	1.0	4.667	19.0	4.222	8.0	4.537	3.0
27-4569 Drama II (S)	4.211	37.0	4.737	11.0	4.579	2.0	4.509	4.0
26-6725 Met. Work, Spec. 5-6 (Y)	4.433	13.0	4.867	3.0	4.067	24.0	4.456	5.0
26-6553 Bookkeeping 3-4 (Y)	4.400	15.0	4.800	5.0	4.000	31.5	4.400	6.0
23-3315 Science 7, ISCS (Y)	4.196	39.0	4.485	37.0	4.456	3.0	4.379	7.0
26-6763 Woodworking 3-4 (Y)	4.536	5.0	4.429	42.5	4.155	14.0	4.373	8.0
26-6729 Vocational Welding (Y)	4.467	12.0	4.733	12.5	3.867	41.0	4.356	9.0
26-5797 Choir, Concert (Y)	4.591	4.0	4.318	61.0	4.114	17.0	4.341	10.0
26-3583 Electronics 3-4 (Y)	4.143	52.5	4.714	14.5	4.143	15.5	4.333	11.5
26-6513 Orna. Horticulture, Adv. (Y-2hrs.)	4.143	52.5	4.714	14.5	4.143	15.5	4.333	11.5
26-6613 Typing 3-4 (Y)	4.533	6.5	4.733	12.5	3.667	70.5	4.311	13.0
23-5331 Home Ec. 8 (Y)	4.240	27.0	4.666	21.0	4.012	27.0	4.306	14.0
26-6565 Business Law (S)	4.043	73.0	4.826	4.0	4.000	31.5	4.290	15.0
22-5377 Chorus 8-9 (Y)	4.323	19.5	4.312	63.0	4.186	9.0	4.274	16.0
26-6715 Adv. Architecture 7-8 (Y)	4.215	33.0	4.643	22.0	3.929	35.0	4.262	17.0
27-6613 Typing 3-4 (Y)	4.175	45.0	4.750	9.5	3.830	45.0	4.252	18.0
26-6767 Vocational Carpentry (Y)	4.188	42.0	4.688	16.5	3.875	40.0	4.250	19.0
22-5303 Art 8-9 (Y)	4.421	14.0	4.158	95.5	4.158	13.0	4.246	20.5
27-6575 Clerical Practice (S)	4.105	60.5	4.579	26.0	4.053	25.0	4.246	20.5
26-6627 Distributive Ed. 12 (Y)	4.176	44.0	4.882	2.0	3.647	76.0	4.235	22.0
23-5411 Physical Education 7 (Y)	4.154	49.0	4.242	78.0	4.302	6.0	4.233	23.0

APPENDIX D

Franklin Pierce Cost-Effective Ratio

Evaluation is an all-encompassing term which takes different meanings in different situations. Constant evaluation takes place as the District progresses toward meeting its objectives. This involves a simple decision on the part of each person assigned a task within the program: "Am I doing what I am supposed to do?" If the program is in sufficient detail, the person who is doing the task, as well as the person or persons responsible for managing the program, will be able to answer "yes" or "no" at any time.

However, in terms of the total PPDES system, evaluation has a different connotation. Here evaluation is the process by which a cost-effective ratio or matrix is developed to enable further planning decisions to be made.

The essential requirement of a cost-effectiveness model is the ability to link student growth data and educational cost data with program activities. Theoretically, every identifiable program activity has associated with it both a cost factor and a student growth factor. The two factors, when known, can be plotted on a simple matrix to indicate the cost-effectiveness of a particular program activity.

For example, the following procedures yield cost-effective information to allow administrators to make decisions about the selection of reading program activities:

1. Enumerate all the reading curricula operating in the District in a given year.
2. Identify the actual instructional groups functioning within each reading curriculum.
3. Identify the students enrolled in each instructional group.
4. Accumulate relevant measures (reading scores on the CTBS, attitudinal and self-concept measures) on the students at two points in time - before the beginning of the curriculum and as close to the end of the instruction as possible.
5. Calculate the mean scores for each

instructional group for each measure in time.

6. Regress post-test means on pre-test means to generate student growth measures for each instructional unit. From this regression equation a predicted post-test mean can be generated. The difference between the predicted and actual post-test means will be the measure of student growth.
7. Group the student growth measures according to the reading programs used. Develop a new mean (the sum of all growth scores in a given curriculum) for each reading program. (These analyses also can be done separately by sex or grade).
8. Compute a cost factor based upon both specific and pro-rated costs. Calculate the mean costs for each reading program.

9. Each reading program can now be located within a matrix in which columns represent levels of student growth and rows represent levels of cost (see the sample matrix below).

The evaluation task can be more detailed for several matrices can be constructed for reading programs, one for each measure of interest. In such multiple matrices, programs in a given curricular area will vary only in location by column, not by row.

Decision makers, when viewing three matrices representing achievement, attitude and self-concept, will need to determine how they want to assign weights to each measure. For example, is the program which generates high student growth in achievement, but low attitudes and self-concept, to be preferred to one which generates relatively low achievement but high attitudes and self-concept?

Cost-Effectiveness Matrix

		Reading							
		STUDENT GROWTH							
		low				high			
PROGRAM ACTIVITY COSTS	low								
	high								

Diagram illustrating the Cost-Effectiveness Matrix. The matrix is a 4x4 grid with columns representing Student Growth (low to high) and rows representing Program Activity Costs (low to high). The matrix is divided into four quadrants labeled A, B, C, and D. Quadrant A (top-left) is labeled 'Least Preferred Programs'. Quadrant B (top-right) is labeled 'Most Preferred Programs'. Quadrant C (bottom-left) is labeled 'Vacuous Programs'. Quadrant D (bottom-right) is labeled 'Hard Decision Programs'. The matrix is also labeled with 'E' and 'F' in the top row, 'G' and 'H' in the bottom row, and 'A' and 'B' in the middle columns.

Appendix E
General Design for Evaluating the Relative
Effectiveness of Basic Skills Programs Within a School District

O.G.
2/27/76

Potential Input Variable Set									Control Variable (Pre Test Mean)	Students	Instructional Groups	Outcome Variable of Interest (Post Test)
Program Identification	Instructional Characteristics	Student Characteristics	Teacher Characteristics	Grade Level	Duration Of Program	Student Time In Program	Teacher/Student Ratio	Other				
									$\bar{x}_{A1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	A	$\bar{x}_{A2}$
									$\bar{x}_{B1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	B	$\bar{x}_{B2}$
									$\bar{x}_{C1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	C	$\bar{x}_{C2}$
									$\bar{x}_{D1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	D	$\bar{x}_{D2}$
									$\bar{x}_{E1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	E	$\bar{x}_{E2}$
									$\bar{x}_{F1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	F	$\bar{x}_{F2}$
									$\vdots$	$\vdots$	$\vdots$	$\vdots$
									$\bar{x}_{Z1}$	$\left\{ \begin{matrix} 1 \\ 2 \\ \dots \\ n \end{matrix} \right\}$	Z	$\bar{x}_{Z2}$

Appendix F
COURSE INFORMATION
ELEMENTARY SCHOOL SOCIAL STUDIES
Franklin Pierce School District, Tacoma, Washington

COURSE NAME AND NUMBER: Social Studies-Grade 1 1000

HISTORY AND DEVELOPMENT: Course was developed prior to ESP and was revised through ESP support.

COURSE DESCRIPTION: Basic text is Harcourt Brace, Concepts and Values with supplementary texts, Allyn & Bacon, Our Country and Explorers and Discoverers. Use of Holt Databank and text, Inquiring About People. Studies physical likenesses and differences among children, adaptation to environment, learning through people interaction, purpose and enforcement of rules, human variability in physical and emotional needs, economic use of natural and human resources, and historical relationships and planning. Uses Nystrom Readiness Wall Map of U.S. This course is offered at Brookdale, Christensen, Collins, Elmhurst, Harvard, and James Sales Elementary Schools. Typical class size is 26 ranging from 17 to 31.

INSTRUCTIONAL CHARACTERISTICS: Typical characteristics of instruction are checked below for each teacher of this course. Teacher code numbers with names are found on the last pages of Course Information.

TEACHER CODES							INSTRUCTIONAL CHARACTERISTICS
105	502	524	515	402	401	317	STUDENT PARTICIPATION/INDIVIDUALIZATION
			X			X	Teacher and/or student developed curriculum & instructional materials
X			X				Written course objectives
			X				Students participate in selecting learning objectives
X			X				Pretests (diagnostic and/or placement)
				X	X	X	Student progresses at her/his own rate
	X			X			All students pursue same objectives, use same means
X			X		X	X	All students pursue same objectives, use varied means
							All students pursue varied objectives, use varied means
			X				Records kept of student progress on course objectives
X			X				Post tests
							Student advancement or credit based upon completion of objectives
							PRESENTATION
X	X	X		X	X		Teacher presentation/lecture
	X	X	X	X	X	X	Discovery techniques; questioning strategies
X	X	X	X	X	X	X	Large group instruction
						X	Team teaching
X	X			X		X	Small group instruction
							INSTRUCTIONAL ASSISTANTS
							Paid teacher aides
							Adult volunteer teacher aides
							Peer or cross-age tutors
						X	Student assistants
							LEARNING RESOURCES
						X	Single textbook
X		X	X		X	X	Multiple textbooks
X			X	X		X	Supplemental workbook or worksheets
X	X	X	X	X	X	X	Individual and/or group projects
X		X		X	X	X	Out-of-school activities (e.g., field trips, work experience)
X	X	X	X	X	X	X	Audio-visual techniques
							Self-instruction packets (LAPS, etc.)
			X		X	X	Classroom learning centers
			X	X		X	Library and/or resource center
						X	Student use of instructional machines
							SKILL MAINTENANCE
							Drill and practice in basic facts
X		X	X		X	X	Teacher-assigned seatwork
							Homework